

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2004, 16:46:52 ; Search time 144 Seconds
(without alignments)
63.940 Million cell updates/sec

Title: US-10-783-095-10

Perfect score: 131

Sequence: 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications 2A:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	131	100.0	54	9	US-09-815-229-13
2	131	100.0	54	16	US-10-740-098-13
3	131	100.0	86	9	US-09-878-380-1
4	131	100.0	86	10	US-09-858-935B-4
5	131	100.0	86	13	US-10-028-410-2
6	131	100.0	86	13	US-10-054-873-4
7	131	100.0	86	14	US-10-444-326-2
8	131	100.0	86	14	US-10-271-869-4
9	131	100.0	86	15	US-10-444-262-2
10	131	100.0	86	15	US-10-444-649-2
11	131	100.0	86	15	US-10-444-701-2
12	131	100.0	96	9	US-09-947-563-4
13	131	100.0	96	9	US-09-947-563-5

14	131	100.0	110	9	US-09-205-658-125	Sequence 125, Appl
15	131	100.0	110	9	US-09-815-229-3	Sequence 3, Appli
16	131	100.0	110	9	US-09-804-409A-9	Sequence 9, Appli
17	131	100.0	110	10	US-09-969-748C-6	Sequence 6, Appli
18	131	100.0	110	10	US-09-963-893-125	Sequence 125, App
19	131	100.0	110	14	US-10-038-686-1	Sequence 1, Appli
20	131	100.0	110	14	US-10-328-813-2	Sequence 2, Appli
21	131	100.0	110	14	US-10-383-285-2	Sequence 2, Appli
22	131	100.0	110	14	US-10-346-563-2	Sequence 2, Appli
23	131	100.0	110	15	US-10-321-717-2	Sequence 2, Appli
24	131	100.0	110	15	US-10-411-037-44	Sequence 44, Appl
25	131	100.0	110	15	US-10-411-026-44	Sequence 44, Appl
26	131	100.0	110	15	US-10-410-962-44	Sequence 44, Appl
27	131	100.0	110	15	US-10-411-049-44	Sequence 44, Appl
28	131	100.0	110	15	US-10-700-725-20	Sequence 20, Appl
29	131	100.0	110	16	US-10-410-930-44	Sequence 44, Appl
30	131	100.0	110	16	US-10-410-997-44	Sequence 44, Appl
31	131	100.0	110	16	US-10-411-012-44	Sequence 44, Appl
32	131	100.0	110	16	US-10-287-994-44	Sequence 44, Appl
33	131	100.0	110	16	US-10-740-098-3	Sequence 3, Appli
34	131	100.0	110	16	US-10-410-913-44	Sequence 44, Appl
35	131	100.0	117	9	US-09-280-030-63	Sequence 63, Appl
36	131	100.0	130	9	US-09-280-030-62	Sequence 62, Appl
37	128	97.7	100	17	US-10-425-115-337310	Sequence 337310,
38	125	95.4	110	16	US-10-419-539-5	Sequence 5, Appli
39	125	95.4	124	15	US-10-221-677-24	Sequence 24, Appl
40	107	81.7	33	17	US-10-641-834-24	Sequence 24, Appl
41	107	81.7	35	9	US-09-815-229-4	Sequence 4, Appli
42	107	81.7	35	9	US-09-947-563-3	Sequence 3, Appli
43	107	81.7	35	16	US-10-740-098-4	Sequence 4, Appli
44	100	76.3	33	9	US-09-815-229-9	Sequence 9, Appli
45	100	76.3	33	16	US-10-740-098-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-815-229-13
; Sequence 13, Application US/09815229
; Patent No. US20020058614A1
; GENERAL INFORMATION:
; APPLICANT: Filvaroff, Ellen H.
; APPLICANT: Okumu, Franklin W.
; TITLE OF INVENTION: USE OF INSULIN FOR THE TREATMENT OF CARTILAGINOUS DISORDERS
; FILE REFERENCE: F1786RIUS
; CURRENT APPLICATION NUMBER: US/09/815,229
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/192,103
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 13
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Portion of insulin variant from EP 171,147.
US-09-815-229-13

Query Match 100.0%; Score 131; DB 9; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26

Db 13 GGGPGAGSLQPLALEGSLQKRGIVEQ 38

RESULT 2

US-10-740-098-13
; Sequence 13, Application US/10740098
; Publication No. US20040138101A1
; GENERAL INFORMATION:

```

; APPLICANT: Filvaroff, Ellen H.
; APPLICANT: Okumu, Franklin W.
; TITLE OF INVENTION: USE OF INSULIN FOR THE TREATMENT OF CARTILAGINOUS DISORDERS
; FILE REFERENCE: P1786RLUS
; CURRENT APPLICATION NUMBER: US/10/740,098
; CURRENT FILING DATE: 2003-12-17
; PRIOR APPLICATION NUMBER: US/09/815,229
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/192,103
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 13
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Portion of insulin variant from EP 171,147.
US-10-740-098-13

```

```

Query Match      100.0%; Score 131; DB 16; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
    |||||
DB 13 GGGPGAGSLQPLALEGSLQKRGIVEQ 38
    |||||

```

```

RESULT 3
US-09-878-380-1
; Sequence 1, Application US/09878380
; Patent No. US20020160435A1
; GENERAL INFORMATION:
; APPLICANT: Fujirebio Inc.
; APPLICANT: KITAJIMA, Sachiko
; APPLICANT: KURANO, Yoshihiro
; APPLICANT: NAKATSUBO, Kaoru
; APPLICANT: NISHIZONO, Isao
; TITLE OF INVENTION: Immunoassay For Measuring Human C-Peptide and Kit Therefor
; FILE REFERENCE: 0760-0291P
; CURRENT APPLICATION NUMBER: US/09/878,380
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: JP 2000-174691
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-878-380-1

```

```

Query Match      100.0%; Score 131; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
    |||||
DB 45 GGGPGAGSLQPLALEGSLQKRGIVEQ 70
    |||||

```

```

RESULT 4
US-09-858-935B-4
; Sequence 4, Application US/09858935B
; Publication No. US20030069177A1
; GENERAL INFORMATION:
; APPLICANT: Dubaquié, Yves
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Lowman, Henry B.
; TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
; FILE REFERENCE: P1794R1
; CURRENT APPLICATION NUMBER: US/09/858,935B
; CURRENT FILING DATE: 2002-07-02

```

```

; PRIOR APPLICATION NUMBER: US 60/248,985
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 60/204,490
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 4
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-935B-4

```

```

Query Match      100.0%; Score 131; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
    |||||
DB 45 GGGPGAGSLQPLALEGSLQKRGIVEQ 70
    |||||

```

```

RESULT 5
US-10-028-410-2
; Sequence 2, Application US/10028410
; Publication No. US20020160955A1
; GENERAL INFORMATION:
; APPLICANT: Dubaquié, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1-1
; CURRENT APPLICATION NUMBER: US/10/028,410
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US/09/477,924
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-410-2

```

```

Query Match      100.0%; Score 131; DB 13; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
    |||||
DB 45 GGGPGAGSLQPLALEGSLQKRGIVEQ 70
    |||||

```

```

RESULT 6
US-10-054-873-4
; Sequence 4, Application US/10054873
; Publication No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ru
; TITLE OF INVENTION: Chimeric Protein Containing an
;                               Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 22, 2004, 20:14:29 ; Search time 339 Seconds
(without alignments)
414.437 Million cell updates/sec

Title: US-10-783-095-10

Perfect score: 131

Sequence: 1 GGPGAGSLOPAALEGSQKRGIVEQ 26

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Xgapop 10.0, Xgapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10783095/runat_22112004_143538_25412/app_query.fasta_1.199
-DB=Published Applications NA -CFMT=fastap -SUFFIX=P2N.rnpb -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10783095 @CGN 1 1 575 @runat_22112004_143538_25412
-NCPV=6 -ICPV=3 -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

§

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	131	100.0	258	15	US-10-383-285-1
2	131	100.0	260	10	US-09-807-742-16
3	131	100.0	260	10	US-09-807-742-17
4	131	100.0	260	17	US-10-746-149-23
5	131	100.0	260	18	US-10-609-019-21
6	131	100.0	330	15	US-10-346-563-1
7	131	100.0	330	16	US-10-321-717-1
8	131	100.0	351	9	US-09-280-030-49
9	131	100.0	390	9	US-09-280-030-48
10	131	100.0	450	9	US-09-804-409A-10
11	131	100.0	450	10	US-09-969-748C-5
12	131	100.0	450	15	US-10-328-813-1
13	131	100.0	450	16	US-10-411-037-43
14	131	100.0	450	16	US-10-411-026-43
15	131	100.0	450	16	US-10-410-962-43
16	131	100.0	450	16	US-10-411-049-43
17	131	100.0	450	17	US-10-410-930-43
18	131	100.0	450	17	US-10-410-997-43
19	131	100.0	450	17	US-10-411-012-43
20	131	100.0	450	17	US-10-287-994-43
21	131	100.0	450	17	US-10-410-913-43
22	131	100.0	1182	9	US-09-919-344-27
23	131	100.0	1826	18	US-10-609-019-27
24	131	100.0	2499	10	US-09-775-508A-1
25	131	100.0	10487	17	US-10-746-149-44
26	131	100.0	10487	18	US-10-609-019-32
27	131	100.0	10512	17	US-10-746-149-42
28	131	100.0	10512	18	US-10-609-019-31
29	131	100.0	10880	17	US-10-746-149-47
30	131	100.0	10895	18	US-10-609-019-41
31	131	100.0	11255	17	US-10-746-149-43
32	131	100.0	11271	18	US-10-609-019-42
33	131	100.0	11332	18	US-10-609-019-43
34	128	97.7	348	18	US-10-425-115-152647
35	125	95.4	333	17	US-10-419-539-2
36	125	95.4	375	16	US-10-221-677-25
37	125	95.4	598	13	US-10-013-032-11
38	125	95.4	633	13	US-10-013-032-12
39	125	95.4	666	13	US-10-013-032-13
40	125	95.4	1080	17	US-10-419-539-3
41	125	95.4	1217	17	US-10-419-539-4
42	125	95.4	1944	16	US-10-221-677-26
43	111	84.7	1498	9	US-09-925-297-345
44	99	75.6	4646	16	US-10-430-752A-14
45	98	74.8	390	9	US-09-833-381-939

ALIGNMENTS

RESULT 1

US-10-383-285-1
; Sequence 1, Application US/10383285
; Publication No. US20030200566A1
; GENERAL INFORMATION:
; APPLICANT: STREATHFIELD, STEPHEN
; APPLICANT: HOWARD, JOHN
; TITLE OF INVENTION: PRODUCTION OF INSULIN AND INSULIN-LIKE PROTEINS IN PLANTS
; FILE REFERENCE: 10040
; CURRENT APPLICATION NUMBER: US/10/383,285
; CURRENT FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: 60/362,874
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 258
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-383-285-1

Alignment Scores:
Pred. No.: 4,73e-11 Length: 258
Score: 131.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-783-095-10 (1-26) x US-10-383-285-1 (1-258)

Qy 1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGluGlySerLeuGlnLys 20
Db 133 GCGGTGGCCCGGGCGCGGCTCCCTCCAGCGCTCGCGCTCGAGGGCTCCCTCCAGAAG 192
Qy 21 ArgGlyIleValGluGln 26
Db 193 CGCGGTATGCTGTGACAA 210

RESULT 2

US-09-807-742-16

; Sequence 16, Application US/09807742

; Publication No. US20030204864A1

; GENERAL INFORMATION:

; APPLICANT: DANIELL, HENRY

; TITLE OF INVENTION: PRODUCTION OF PHARMACEUTICAL PROTEINS IN TRANSGENIC

; FILE REFERENCE: 1465-PCT-US-00

; CURRENT APPLICATION NUMBER: US/09/807,742

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: PCT/US01/06288

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 16

; LENGTH: 260

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-807-742-16

Alignment Scores:
Pred. No.: 4,73e-11 Length: 260
Score: 131.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-783-095-10 (1-26) x US-09-807-742-16 (1-260)

Qy 1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGluGlySerLeuGlnLys 20
Db 133 GCGGGGGCCCTGGTGGCAGGAGCGCTGAGCCCTGGCGCTGGAGGGTCCCTCAGAAG 192
Qy 21 ArgGlyIleValGluGln 26
Db 193 CGTGGCATTTGTGACAA 210

RESULT 3

US-09-807-742-17

; Sequence 17, Application US/09807742

; Publication No. US20030204864A1

; GENERAL INFORMATION:

; APPLICANT: DANIELL, HENRY

; TITLE OF INVENTION: PRODUCTION OF PHARMACEUTICAL PROTEINS IN TRANSGENIC

; FILE REFERENCE: 1465-PCT-US-00

; CURRENT APPLICATION NUMBER: US/09/807,742

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: PCT/US01/06288

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 17
; LENGTH: 260
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chloroplast
; OTHER INFORMATION: modified proinsulin sequence
US-09-807-742-17

Alignment Scores:

Pred. No.: 4,73e-11 Length: 260
Score: 131.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-783-095-10 (1-26) x US-09-807-742-17 (1-260)

Qy 1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGluGlySerLeuGlnLys 20
Db 133 GGTGGTGGTCTGGTGGTCTTTTACACCTTTAGCTTTAGAGGTTCTTTACAAAA 192
Qy 21 ArgGlyIleValGluGln 26
Db 193 CGTGGTATTGTGACAA 210

RESULT 4

US-10-746-149-23

; Sequence 23, Application US/10746149

; Publication No. US20040172867A1

; GENERAL INFORMATION:

; APPLICANT: Cooper, Richard K.

; APPLICANT: Fioretti, William C.

; APPLICANT: Cadd, Gary G.

; TITLE OF INVENTION: Administration of Transposon-Based Vectors to Reproductive Organs

; FILE REFERENCE: 51687-0280 (51687-294923)

; CURRENT APPLICATION NUMBER: US/10/746,149

; CURRENT FILING DATE: 2003-12-24

; PRIOR APPLICATION NUMBER: US 60/392,415

; PRIOR FILING DATE: 2002-06-26

; PRIOR APPLICATION NUMBER: US 60/441,381

; PRIOR FILING DATE: 2003-01-21

; PRIOR APPLICATION NUMBER: US 60/441,447

; PRIOR FILING DATE: 2003-01-21

; PRIOR APPLICATION NUMBER: US 60/441,405

; PRIOR FILING DATE: 2003-01-21

; PRIOR APPLICATION NUMBER: US 60/441,502

; PRIOR FILING DATE: 2003-01-21

; PRIOR APPLICATION NUMBER: US 60/441,392

; PRIOR FILING DATE: 2003-01-21

; PRIOR APPLICATION NUMBER: US 60/441,377

; PRIOR FILING DATE: 2003-01-21

; PRIOR APPLICATION NUMBER: US 10/609,019

; PRIOR FILING DATE: 2003-06-26

; NUMBER OF SEQ ID NOS: 52

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 23

; LENGTH: 260

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

US-10-746-149-23

Alignment Scores:

Pred. No.: 4,73e-11 Length: 260
Score: 131.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

Result No.	Score	Query Match	Length	DB	ID	Description
1	26	100.0	54	4	US-09-815-229-13	Sequence 13, Appl
2	26	100.0	86	4	US-09-477-924-2	Sequence 2, Appl
3	26	100.0	86	4	US-09-723-981-2	Sequence 2, Appl
4	26	100.0	86	4	US-09-723-896-2	Sequence 2, Appl
5	26	100.0	86	4	US-09-878-380-1	Sequence 1, Appl
6	26	100.0	90	1	US-08-030-731A-43	Sequence 43, Appl
7	26	100.0	91	4	US-09-676-787-7	Sequence 7, Appl
8	26	100.0	96	2	US-09-134-836-4	Sequence 4, Appl
9	26	100.0	96	2	US-09-134-835-5	Sequence 5, Appl
10	26	100.0	96	3	US-09-386-303A-4	Sequence 4, Appl
11	26	100.0	96	3	US-09-386-303A-5	Sequence 5, Appl
12	26	100.0	96	4	US-09-947-563-4	Sequence 4, Appl
13	26	100.0	96	4	US-09-947-563-5	Sequence 5, Appl
14	26	100.0	97	1	US-08-160-376A-4	Sequence 4, Appl
15	26	100.0	97	1	US-08-389-487-7	Sequence 7, Appl
16	26	100.0	97	3	US-09-099-307-6	Sequence 6, Appl
17	26	100.0	97	3	US-09-099-307-7	Sequence 7, Appl
18	26	100.0	97	3	US-09-099-307-8	Sequence 8, Appl
19	26	100.0	97	3	US-09-099-307-11	Sequence 11, Appl
20	26	100.0	98	4	US-09-701-968-7	Sequence 7, Appl
21	26	100.0	99	4	US-09-701-968-8	Sequence 8, Appl
22	26	100.0	100	4	US-09-701-968-9	Sequence 9, Appl
23	26	100.0	110	3	US-08-950-720A-11	Sequence 11, Appl
24	26	100.0	110	3	US-08-589-028-2	Sequence 2, Appl
25	26	100.0	110	3	US-08-589-028-4	Sequence 4, Appl
26	26	100.0	110	3	US-08-784-583-2	Sequence 2, Appl
27	26	100.0	110	3	US-08-784-583-4	Sequence 4, Appl

Best Local Similarity 100.0%; Pred. No. 5.7e-18; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
Db 45 GGGPGAGSLQPLALEGSLQKRGIVEQ 70

RESULT 3

US-09-723-981-2
; Sequence 2, Application US/097233981
; Patent No. 6506874
; GENERAL INFORMATION:
; APPLICANT: Dubaquis, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/09/723,981
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-723-981-2

Query Match 100.0%; Score 26; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 5.7e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
Db 45 GGGPGAGSLQPLALEGSLQKRGIVEQ 70

RESULT 4

US-09-723-896-2
; Sequence 2, Application US/09723896
; Patent No. 6509443
; GENERAL INFORMATION:
; APPLICANT: Dubaquis, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/09/723,896
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-723-896-2

Query Match 100.0%; Score 26; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 5.7e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
Db 45 GGGPGAGSLQPLALEGSLQKRGIVEQ 70

RESULT 5

US-09-878-380-1
; Sequence 1, Application US/09878380
; Patent No. 6534281
; GENERAL INFORMATION:
; APPLICANT: Fujirebio Inc.
; APPLICANT: KITAJIMA, Sachiko

; APPLICANT: KURANO, Yoshihiro
; APPLICANT: NAKATSUBO, Kaoru
; APPLICANT: NISHIZONO, Isao
; TITLE OF INVENTION: Immunoassay For Measuring Human C-Peptide and Kit Therefor
; FILE REFERENCE: 0760-0291P
; CURRENT APPLICATION NUMBER: US/09/878,380
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: JP 2000-174691
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-878-380-1

Query Match 100.0%; Score 26; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 5.7e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
Db 45 GGGPGAGSLQPLALEGSLQKRGIVEQ 70

RESULT 6

US-08-030-731A-43
; Sequence 43, Application US/08030731A
; Patent No. 5426036
; GENERAL INFORMATION:
; APPLICANT: Koller, Klaus-Peter
; APPLICANT: Riess, Guenther Johannes
; APPLICANT: Uhlmann, Eugen
; APPLICANT: Wallmeier, Holger
; TITLE OF INVENTION: Processes for the Preparation of Foreign
; TITLE OF INVENTION: Proteins in Streptomyces
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/030,731A
; FILING DATE: 12-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/189,840
; FILING DATE: 03-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/430,622
; FILING DATE: 01-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/687,610
; FILING DATE: 19-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/735,757
; FILING DATE: 29-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DB P 37 14 866.4
; FILING DATE: 05-MAY-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DB P 38 37 273.8
; FILING DATE: 03-NOV-1988

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2004, 17:07:53 ; Search time 140 Seconds
(without alignments)
65.767 Million cell updates/sec

Title: US-10-783-095-10
Perfect score: 26
Sequence: 1 GGPGAGSLQPLALEGSLQKRGIVEQ 26

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 1570615 seqs, 354127592 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	54	9	US-09-815-229-13
2	26	100.0	54	16	US-10-740-098-13
3	26	100.0	86	9	US-09-878-380-1
4	26	100.0	86	10	US-09-858-935B-4
5	26	100.0	86	13	US-10-028-410-2
6	26	100.0	86	13	US-10-054-873-4
7	26	100.0	86	14	US-10-444-326-2
8	26	100.0	86	14	US-10-271-869-4
9	26	100.0	86	15	US-10-444-262-2
10	26	100.0	86	15	US-10-444-649-2
11	26	100.0	86	15	US-10-444-701-2
12	26	100.0	96	9	US-09-947-563-4
13	26	100.0	96	9	US-09-947-563-5

14	26	100.0	110	9	US-09-205-658-125	Sequence 125, Appl
15	26	100.0	110	9	US-09-815-229-3	Sequence 3, Appl
16	26	100.0	110	9	US-09-804-409A-9	Sequence 9, Appl
17	26	100.0	110	10	US-09-969-748C-6	Sequence 6, Appl
18	26	100.0	110	10	US-09-963-693-125	Sequence 125, Appl
19	26	100.0	110	14	US-10-038-686-1	Sequence 1, Appl
20	26	100.0	110	14	US-10-328-813-2	Sequence 2, Appl
21	26	100.0	110	14	US-10-383-285-2	Sequence 2, Appl
22	26	100.0	110	14	US-10-346-563-2	Sequence 2, Appl
23	26	100.0	110	15	US-10-321-717-2	Sequence 2, Appl
24	26	100.0	110	15	US-10-411-037-44	Sequence 44, Appl
25	26	100.0	110	15	US-10-411-026-44	Sequence 44, Appl
26	26	100.0	110	15	US-10-410-962-44	Sequence 44, Appl
27	26	100.0	110	15	US-10-411-049-44	Sequence 44, Appl
28	26	100.0	110	15	US-10-700-725-20	Sequence 20, Appl
29	26	100.0	110	16	US-10-410-930-44	Sequence 44, Appl
30	26	100.0	110	16	US-10-410-997-44	Sequence 44, Appl
31	26	100.0	110	16	US-10-411-012-44	Sequence 44, Appl
32	26	100.0	110	16	US-10-287-994-44	Sequence 44, Appl
33	26	100.0	110	16	US-10-740-098-3	Sequence 3, Appl
34	26	100.0	110	16	US-10-410-913-44	Sequence 44, Appl
35	26	100.0	117	9	US-09-280-030-63	Sequence 63, Appl
36	26	100.0	130	9	US-09-280-030-62	Sequence 62, Appl
37	24	92.3	100	17	US-10-425-115-337310	Sequence 337310,
38	21	80.8	33	17	US-10-641-834-24	Sequence 24, Appl
39	21	80.8	35	9	US-09-815-229-4	Sequence 4, Appl
40	21	80.8	35	9	US-09-947-563-3	Sequence 3, Appl
41	21	80.8	35	16	US-10-740-098-4	Sequence 4, Appl
42	19	73.1	31	9	US-09-269-439-1	Sequence 1, Appl
43	19	73.1	31	9	US-09-878-380-2	Sequence 2, Appl
44	19	73.1	31	15	US-10-430-752A-1	Sequence 1, Appl
45	19	73.1	31	15	US-10-430-752A-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-09-815-229-13
; Sequence 13, Application US/09815229
; Patent No. US20020058614A1
; GENERAL INFORMATION:
; APPLICANT: Filvaroff, Ellen H.
; APPLICANT: Okumu, Franklin W.
; FILE OF INVENTION: USE OF INSULIN FOR THE TREATMENT OF CARTILAGINOUS DISORDERS
; TITLE REFERENCE: P1786RIUS
; CURRENT APPLICATION NUMBER: US/09/815,229
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/192,103
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 13
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Portion of insulin variant from EP 171,147.
US-09-815-229-13

Query Match 100.0%; Score 26; DB 9; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.9e-16;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGAGSLQPLALEGSLQKRGIVEQ 26
Db 13 GGPGAGSLQPLALEGSLQKRGIVEQ 38

RESULT 2

US-10-740-098-13
; Sequence 13, Application US/10740098
; Publication No. US20040138101A1
; GENERAL INFORMATION:

APPLICANT: Filvaroff, Ellen H.
APPLICANT: Okumu, Franklin W.
TITLE OF INVENTION: USE OF INSULIN FOR THE TREATMENT OF CARTILAGENOUS DISORDERS
FILE REFERENCE: P1786R1US
CURRENT APPLICATION NUMBER: US/10/740,098
CURRENT FILING DATE: 2003-12-17
PRIOR APPLICATION NUMBER: US/09/815,229
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: US 60/192,103
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 17
SEQ ID NO 13
LENGTH: 54
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Portion of insulin variant from EP 171,147.
US-10-740-098-13

Query Match 100.0%; Score 26; DB 16; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.9e-16;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
Db 13 GGGPGAGSLQPLALEGSLQKRGIVEQ 38

RESULT 3
US-09-878-380-1
Sequence 1, Application US/09878380
Patent No. US20020160435A1
GENERAL INFORMATION:
APPLICANT: Fujirebio Inc.
APPLICANT: KITAJIMA, Sachiko
APPLICANT: KURANO, Yoshihiro
APPLICANT: NAKATSUBO, Kaoru
APPLICANT: NISHIZONO, Isao
TITLE OF INVENTION: Immunoassay For Measuring Human C-Peptide and Kit Therefor
FILE REFERENCE: 0760-0291P
CURRENT APPLICATION NUMBER: US/09/878,380
CURRENT FILING DATE: 2001-06-12
PRIOR APPLICATION NUMBER: JP 2000-174691
PRIOR FILING DATE: 2000-06-12
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
US-09-878-380-1

Query Match 100.0%; Score 26; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
Db 45 GGGPGAGSLQPLALEGSLQKRGIVEQ 70

RESULT 4
US-09-858-935B-4
Sequence 4, Application US/09858935B
Publication No. US20030069177A1
GENERAL INFORMATION:
APPLICANT: Dubaigue, Yves
APPLICANT: Filvaroff, Ellen
APPLICANT: Lowman, Henry B.
TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
FILE REFERENCE: P1794R1
CURRENT APPLICATION NUMBER: US/09/858,935B
CURRENT FILING DATE: 2002-07-02

PRIOR APPLICATION NUMBER: US 60/248,985
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 60/204,490
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 153
SEQ ID NO 4
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
US-09-858-935B-4

Query Match 100.0%; Score 26; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
Db 45 GGGPGAGSLQPLALEGSLQKRGIVEQ 70

RESULT 5
US-10-028-410-2
Sequence 2, Application US/10028410
Publication No. US20020160955A1
GENERAL INFORMATION:
APPLICANT: Dubaigue, Yves
APPLICANT: Lowman, Henry
TITLE OF INVENTION: PROTEIN VARIANTS
FILE REFERENCE: P1712R1-1
CURRENT APPLICATION NUMBER: US/10/028,410
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US/09/477,924
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 2
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
US-10-028-410-2

Query Match 100.0%; Score 26; DB 13; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
Db 45 GGGPGAGSLQPLALEGSLQKRGIVEQ 70

RESULT 6
US-10-054-873-4
Sequence 4, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2004, 16:50:52 ; Search time 193 Seconds
(without alignments)

77.512 Million cell updates/sec

Title: US-10-783-095-10

Perfect score: 26

Sequence: 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	65	2 Q8H280	Q8H280 pongo pygma
2	26	100.0	65	2 Q8H281	Q8H281 gorilla gor
3	26	100.0	110	1 INS_CERAE	P30407 cercopithec
4	26	100.0	110	1 INS_HUMAN	P10308 homo sapien
5	26	100.0	110	1 INS_MACFA	P30406 macaca fasc
6	26	100.0	110	1 INS_PANTR	P30410 pan troglod
7	26	100.0	110	1 INS_PONPY	Q8H280 pongo pygma
8	26	100.0	110	2 Q6YK33	Q6YK33 gorilla gor
9	26	100.0	110	2 AAP35454	AAP35454 homo sapi
10	26	100.0	110	2 AAN06935	AAN06935 gorilla g
11	26	100.0	110	2 AAN39451	AAN39451 homo sapi
12	9	34.6	110	1 INS_CANFA	P01321 canis famil
13	9	34.6	110	1 INS_RABIT	P01311 coryctolagus
14	9	34.6	110	1 INS_SPETR	Q91X13 spermophilu
15	9	34.6	110	2 Q8WNW6	Q8WNW6 felis silve
16	8	30.8	105	1 INS_BOVIN	P01317 bos taurus
17	8	30.8	105	1 INS_SHEEP	P01318 ovis aries
18	8	30.8	108	1 INS_PIG	P01315 sus scrofa
19	8	30.8	108	2 AAQ00952	AAQ00952 sus scrof
20	8	30.8	108	2 AAQ00954	AAQ00954 sus scrof
21	8	30.8	108	2 AAQ00957	AAQ00957 sus scrof
22	8	30.8	108	2 AAQ00960	AAQ00960 sus scrof
23	8	30.8	108	2 AAQ00963	AAQ00963 sus scrof
24	8	30.8	108	2 AAQ00966	AAQ00966 sus scrof
25	8	30.8	108	2 AAQ00969	AAQ00969 sus scrof
26	8	30.8	108	2 AAQ00972	AAQ00972 sus scrof
27	8	30.8	108	2 AAQ00975	AAQ00975 sus scrof
28	8	30.8	108	2 AAQ00978	AAQ00978 sus scrof
29	8	30.8	108	2 AAQ00981	AAQ00981 sus scrof
30	8	30.8	108	2 AAQ00983	AAQ00983 sus scrof
31	8	30.8	108	2 AAQ00985	AAQ00985 sus scrof

32 8 30.8 108 2 AAQ00987 AaQ00987 sus scrof
33 8 30.8 108 2 AAQ00990 AaQ00990 sus scrof
34 8 30.8 110 1 INS_PSAOB Q62587 psammomys o
35 8 30.8 447 2 Q9SIA8 Q9SIA8 arabidopsis
36 8 30.8 465 1 FXD3_MOUSE Q61060 mus musculu
37 8 30.8 478 1 FXD3_HUMAN Q9UJ05 homo sapien
38 8 30.8 485 2 Q8PNW5 Q8PNW5 xanthomonas
39 8 30.8 549 2 Q8Z9A9 Q8Z9A9 streptomyce
40 8 30.8 605 2 Q6DI71 Q6DI71 mus musculu
41 8 30.8 631 1 DMK_MOUSE P54265 mus musculu
42 7 26.9 12 2 Q6QHB8 Q6QHB8 homo sapien
43 7 26.9 12 2 AAS55053 Aas55053 homo sapi
44 7 26.9 38 2 Q6XV13 Q6XV13 strix aluco
45 7 26.9 38 2 Q6XVJ2 Q6XVJ2 meleagris g

ALIGNMENTS

RESULT 1

Q8H280 ID Q8H280 PRELIMINARY; PRT; 65 AA.
AC Q8H280;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Insulin (Fragment).
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RA O'huigin C., Tichy H., Klein J.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY092024; AAM76641.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR GO; GO:0007582; P:physiological process; IEA.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
FT NON_TER 1
FT NON_TER 65
SQ SEQUENCE 65 AA; 6920 MW; B772017FDBSCABEA CRC64;

Query Match 100.0%; Score 26; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 4e-17;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
DB 39 GGGPGAGSLQPLALEGSLQKRGIVEQ 64

RESULT 2

Q8H281 ID Q8H281 PRELIMINARY; PRT; 65 AA.
AC Q8H281;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Insulin (Fragment).
OS Gorilla gorilla (Gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA O'huigin C., Tichy H., Klein J.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Belongs to the insulin family.
DR EMBL; AY092023; AAM76640.1; -.

DR GO:0005576; C:extracellular; IEA.
 DR GO:0005179; F:hormone activity; IEA.
 DR GO:0007582; P:physiological process; IEA.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; ILGF; 1.
 DR SMART; SM00078; Insulin; ILGF; 1.
 KW Insulin family.
 FT NON_TER 1
 FT NON_TER 65
 SQ SEQUENCE 65 AA; 6920 MW; B772017FD8BCABEA CRC64;
 Query Match 100.0%; Score 26; DB 2; Length 65;
 Best Local Similarity 100.0%; Pred. No. 4e-17;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
 DB 39 GGGPGAGSLQPLALEGSLQKRGIVEQ 64

RESULT 3
 INS CERAE
 ID INS CERAE STANDARD; PRT; 110 AA.
 AC P0107; P01309;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92219953; PubMed=1560757;
 RA Seino S., Bell G.I., Li W.
 RT "Sequences of primate insulin genes support the hypothesis of a slower
 rate of molecular evolution in humans and apes than in monkeys.";
 RL Mol. Biol. Evol. 9:193-203(1992).
 RN [2]
 RP SEQUENCE OF 57-87.
 RX MEDLINE=72258016; PubMed=4626369;
 RA Peterson J.D., Kehrlich S., Oyer P.E., Steiner D.F.;
 RT "Determination of the amino acid sequence of the monkey, sheep, and
 dog proinsulin C-peptides by a semi-micro Edman degradation
 procedure.";
 RL J. Biol. Chem. 247:4866-4871(1972).
 CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
 increases cell permeability to monosaccharides, amino acids and
 fatty acids. It accelerates glycolysis, the pentose phosphate
 cycle, and glycogen synthesis in liver.
 CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 disulfide bonds.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the insulin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; X61092; CAA43405.1; -
 CC PIR; B42179; B42179.
 CC HSSP; P01308; 1A10.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PRINTS; PR00277; INSULINE.
 DR SMART; SM00078; ILGF; 1.

DR PROSITE; PS00262; INSULIN; 1.
 KW Direct protein sequencing; Glucose metabolism; Hormone;
 KW Insulin family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 Insulin B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin A chain.
 FT DISULFID 31 96 Interchain.
 FT DISULFID 43 109 Interchain.
 FT DISULFID 95 100
 SQ SEQUENCE 110 AA; 12019 MW; 95A1F54BE7B247P9 CRC64;
 Query Match 100.0%; Score 26; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 6.2e-17;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGAGSLQPLALEGSLQKRGIVEQ 26
 DB 69 GGGPGAGSLQPLALEGSLQKRGIVEQ 94

RESULT 4
 INS HUMAN
 ID INS HUMAN STANDARD; PRT; 110 AA.
 AC P01308;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80120725; PubMed=6243748;
 RA Bell G.I., Pictet R.L., Rutter W.J., Cordell B., Tischler E.,
 RA Goodman H.M.;
 RT "Sequence of the human insulin gene.";
 RL Nature 284:26-32(1980).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80236313; PubMed=6248962;
 RA Ullrich A., Dull T.J., Gray A., Brosius J., Sures I.;
 RT "Genetic variation in the human insulin gene.";
 RL Science 209:612-615(1980).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80054779; PubMed=503234;
 RA Bell G.I., Swain W.F., Pictet R.L., Cordell B., Goodman H.M.,
 RA Rutter W.J.;
 RT "Nucleotide sequence of a cDNA clone encoding human preproinsulin.";
 RL Nature 282:525-527(1979).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80147417; PubMed=6927840;
 RA Sures I., Goeddel D.V., Gray A., Ullrich A.;
 RT "Nucleotide sequence of human preproinsulin complementary DNA.";
 RL Science 208:57-59(1980).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93364428; PubMed=8358440;
 RA Lucassen A.M., Bell J.I., Jullier C., Lathrop M.;
 RT "Susceptibility to insulin dependent diabetes mellitus maps to a 4.1
 kb segment of DNA spanning the insulin gene and associated VNTR.";
 RL Nat. Genet. 4:305-310(1993).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Pancreas;
 RC MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2004, 16:58:57 ; Search time 38 Seconds
(without alignments)
65.833 Million cell updates/sec

Title: US-10-783-095-10
Perfect score: 26
Sequence: 1 GGGPGAGSLQALGSLQKRGIVEQ 26

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 79:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	110	1 IPHU	insulin precursor
2	26	100.0	110	2 B42179	insulin precursor
3	26	100.0	110	2 JQ0178	insulin precursor
4	26	100.0	110	2 A42179	insulin precursor
5	9	34.6	110	1 INRB	insulin precursor
6	9	34.6	110	1 IPDG	insulin precursor
7	8	30.8	84	1 IPFG	insulin precursor
8	8	30.8	105	1 IPBO	insulin precursor
9	8	30.8	412	2 I78395	myotonic dystrophy
10	8	30.8	441	2 I78393	myotonic dystrophy
11	8	30.8	447	2 G84687	probable disease x
12	8	30.8	474	2 I78396	myotonic dystrophy
13	8	30.8	516	2 I78394	myotonic dystrophy
14	8	30.8	557	2 S71829	serine/threonine-s
15	7	26.9	103	2 I51221	insulin precursor
16	7	26.9	105	1 IPON	insulin I precursor
17	7	26.9	106	1 IPXL1	insulin I precursor
18	7	26.9	106	1 IPXL2	insulin II precurs
19	7	26.9	107	1 IPCH	insulin precursor
20	7	26.9	108	1 IPCA	insulin precursor
21	7	26.9	109	1 IPRTU	insulin precursor
22	7	26.9	110	1 INMS2	insulin 2 precursor
23	7	26.9	110	1 IPGP	insulin precursor
24	7	26.9	110	1 IPRT2	insulin 2 precursor
25	7	26.9	115	1 IPHF	insulin precursor
26	7	26.9	115	1 IPAF	insulin precursor
27	7	26.9	234	2 G95989	hypothetical glyci
28	7	26.9	245	2 A70777	hypothetical prote
29	7	26.9	277	1 ESPSRK	2,6-dioxo-6-phenyl

RESULT 1

IPHU

insulin precursor [validated] - human

N;Alternate names: preproinsulin

C;Species: Homo sapiens (man)

C;Date: 23-Oct-1981 #sequence revision 23-Oct-1981 #text change 09-Jul-2004

C;Accession: A93222; A94253; A93144; A92075; A91186; I58114; A01579; S581

R;Bell, G.I.; Pictet, R.L.; Rutter, W.J.; Cordell, B.; Tischer, E.; Goodman, H.M.

Nature 284, 26-32, 1980

A;Title: Sequence of the human insulin gene.

A;Reference number: A93222; MUID:80120725; PMID:6243748

A;Accession: A93222

A;Molecule type: DNA

A;Residues: 1-110 <BEL>

A;Cross-references: UNIPROT:P01308; GB:J00265; NID:G186429; PIDN:AAA59172.1; PID:G386828

R;Ullrich, A.; Dull, T.J.; Gray, A.; Brosius, J.; Sures, I.

Science 209, 612-615, 1980

A;Title: Genetic variation in the human insulin gene.

A;Reference number: A94253; MUID:80236313; PMID:6248962

A;Accession: A94253

A;Molecule type: DNA

A;Residues: 1-110 <ULL>

A;Cross-references: GB:J00265; NID:G186429; PIDN:AAA59172.1; PID:G386828

R;Bell, G.I.; Swain, W.F.; Pictet, R.; Cordell, B.; Goodman, H.M.; Rutter, W.J.

Nature 282, 525-527, 1979

A;Title: Nucleotide sequence of a cDNA clone encoding human preproinsulin.

A;Reference number: A93216; MUID:80054779; PMID:503234

A;Accession: A93216

A;Molecule type: mRNA

A;Residues: 1-110 <BEL2>

A;Cross-references: GB:J00265; NID:G186429; PIDN:AAA59172.1; PID:G386828

R;Sures, I.; Goeddel, D.V.; Gray, A.; Ullrich, A.

Science 208, 57-59, 1980

A;Title: Nucleotide sequence of human preproinsulin complementary DNA.

A;Reference number: A94251; MUID:80147417; PMID:6927840

A;Accession: A94251

A;Molecule type: mRNA

A;Residues: 1-110 <SUR>

A;Cross-references: GB:J00265; NID:G186429; PIDN:AAA59172.1; PID:G386828

R;Nicol, D.S.H.W.; Smith, L.F.

Nature 187, 483-485, 1960

A;Title: Amino-acid sequence of human insulin.

A;Reference number: A93144

A;Accession: A93144

A;Molecule type: protein

A;Residues: 25-54;90-110 <NIC>

R;Over, P.E.; Cho, S.; Peterson, J.D.; Steiner, D.F.

J. Biol. Chem. 246, 1375-1386, 1971

A;Title: Studies on human proinsulin. Isolation and amino acid sequence of the human pan

A;Reference number: A92075; MUID:71116410; PMID:5101771

A;Accession: A92075

A;Molecule type: protein

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 22, 2004, 21:32:18 ; Search time 2044 Seconds
(without alignments)
463.519 Million cell updates/sec

Title: US-10-783-095-10

Perfect score: 26

Sequence: 1 GGCGAGSLQPLALEGSLQKRGIVEQ 26

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219665908 residues

Word size: 1

Total number of hits satisfying chosen parameters: 65642373

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2_1/USPTO.spool/US10783095/runat_22112004_143631_25642/app_query.fasta_1.199
-DB=EST -QFWT=fastap -SUFFIX=P2Noligo.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptp
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10783095 @CGN 1 1 3437 @runat_22112004_143631_25642 -NCPU=6 -ICPU=3
-NC MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DGPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST :
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_gsl1.*
9: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	26	100.0	166	4 B1711381	B1711381 ig94h08.x
C 2	26	100.0	168	4 B1711381	B1711381 ig94h08.x
C 3	26	100.0	174	5 BQ272395	BQ272395 ig98c08.x
C 4	26	100.0	179	5 BQ548672	BQ548672 ig98c08.x
C 5	26	100.0	185	4 B1711381	B1711381 ig94h08.x
C 6	26	100.0	189	5 BQ267409	BQ267409 ig94h08.x
C 7	26	100.0	190	4 B1711381	B1711381 ig94h08.x
C 8	26	100.0	190	5 B1711381	B1711381 ig94h08.x
C 9	26	100.0	191	5 BQ270646	BQ270646 ig94h08.x

C 10	26	100.0	191	5 BQ271032	BQ271032 im29c07.x
C 11	26	100.0	191	5 BQ271032	BQ271032 im29c07.x
C 12	26	100.0	191	5 BQ271032	BQ271032 im29c07.x
C 13	26	100.0	194	4 BQ565510	BQ565510 im29c07.x
C 14	26	100.0	198	5 BQ784487	BQ784487 im29c07.x
C 15	26	100.0	203	4 BQ506695	BQ506695 im29c07.x
C 16	26	100.0	220	5 BQ286797	BQ286797 im29c07.x
C 17	26	100.0	220	5 BQ286797	BQ286797 im29c07.x
C 18	26	100.0	220	5 BQ286797	BQ286797 im29c07.x
C 19	26	100.0	221	5 BQ271297	BQ271297 im29c07.x
C 20	26	100.0	221	5 BQ271297	BQ271297 im29c07.x
C 21	26	100.0	225	5 BQ270528	BQ270528 im29c07.x
C 22	26	100.0	225	5 BQ272471	BQ272471 im29c07.x
C 23	26	100.0	225	5 BQ787442	BQ787442 im29c07.x
C 24	26	100.0	226	5 BQ949768	BQ949768 im29c07.x
C 25	26	100.0	230	5 BQ784426	BQ784426 im29c07.x
C 26	26	100.0	231	1 A1473565	A1473565 im29c07.x
C 27	26	100.0	232	5 BQ579821	BQ579821 im29c07.x
C 28	26	100.0	233	5 BQ128108	BQ128108 im29c07.x
C 29	26	100.0	239	5 BQ074135	BQ074135 im29c07.x
C 30	26	100.0	240	5 BQ270525	BQ270525 im29c07.x
C 31	26	100.0	243	5 BQ272179	BQ272179 im29c07.x
C 32	26	100.0	244	6 BQ841431	BQ841431 im29c07.x
C 33	26	100.0	246	4 BQ857652	BQ857652 im29c07.x
C 34	26	100.0	248	6 BQ842532	BQ842532 im29c07.x
C 35	26	100.0	250	6 BQ847713	BQ847713 im29c07.x
C 36	26	100.0	252	5 BQ271002	BQ271002 im29c07.x
C 37	26	100.0	254	4 BQ310312	BQ310312 im29c07.x
C 38	26	100.0	254	5 BQ787576	BQ787576 im29c07.x
C 39	26	100.0	254	5 BQ948073	BQ948073 im29c07.x
C 40	26	100.0	255	4 B1499525	B1499525 im29c07.x
C 41	26	100.0	259	2 BQ583461	BQ583461 im29c07.x
C 42	26	100.0	259	2 BQ583475	BQ583475 im29c07.x
C 43	26	100.0	259	2 BQ583475	BQ583475 im29c07.x
C 44	26	100.0	261	4 B1499559	B1499559 im29c07.x
C 45	26	100.0	261	5 BQ076352	BQ076352 im29c07.x

ALIGNMENTS

RESULT 1
B1711381/c
LOCUS
DEFINITION

B1711381 166 bp mRNA linear EST 11-MAR-2002
ig94h08.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:5023718
3' similar to SW:INS_HUMAN P01308 INSULIN PRECURSOR. [1] ; mRNA
sequence.

ACCESSION B1711381
VERSION B1711381.1 GI:15687076

KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

REFERENCE 1 (bases 1 to 166)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,

Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blaisstein,A.,
Schmitt,A., Theisinger,B., Ritter,E., Konko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisshvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)

TITLE JOURNAL
COMMENT

Other ESTs: ig94h08.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biochem.harvard.edu
Library was constructed by Dr. J. Ferrer In vivo mass-excised to
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington

University Genome Sequencing Center For information on obtaining a
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40UP from Gibco.
Location/Qualifiers

FEATURES

source
1..166
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5023718"
/tissue_type="insulinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Human insulinoma"
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1:
XhoI; Site 2: EcoRI; Constructed with lambda ZapII system
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
pBluescript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."

ORIGIN

Alignment Scores:
Pred. No.: 6,89e-15 Length: 166
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-783-095-10 (1-26) x BI711381 (1-166)

QY 1 GYGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGlyGlySerLeuGlnLys 20
DB 122 GCGGGGGCCCTGGTGCAGGAGCCTGCAGCCCTGGCCCTGGAGGGTCCCTGCAGAAG 63

QY 21 ArgGlyIleValGluGln 26

DB 62 CGTGCATTGTGGAACAA 45

RESULT 2
BM511718 168 bp mRNA linear EST 15-FEB-2002
LOCUS iJ52d11.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:5634500
DEFINITION 3', similar to SW:INS_HUMAN P01308 INSULIN PRECURSOR. [1] ; mRNA
sequence.

ACCESSION BM511718

VERSION BM511718.1 GI:18682861

KEYWORDS EST.

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 168)

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,

Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,

Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,

Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,

Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R.,

Williams,T., Jackson,Y. and Bowers,Y.

Endocrine Pancreas Consortium

Unpublished (2000)

CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohpc.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
University Genome Sequencing Center For information on obtaining a
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES

Location/Qualifiers
1..168
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5634500"
/tissue_type="insulinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Human insulinoma"
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1:
XhoI; Site 2: EcoRI; Constructed with lambda ZapII system
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
pBluescript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."

ORIGIN

Alignment Scores:
Pred. No.: 6,96e-15 Length: 168
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-783-095-10 (1-26) x BM511718 (1-168)

QY 1 GYGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGlyGlySerLeuGlnLys 20
DB 124 GCGGGGGCCCTGGTGCAGGAGCCTGCAGCCCTGGCCCTGGAGGGTCCCTGCAGAAG 65

QY 21 ArgGlyIleValGluGln 26

DB 64 CGTGCATTGTGGAACAA 47

RESULT 3

BQ272395/c

LOCUS BQ272395.1

DEFINITION 3', similar to SW:INS_HUMAN P01308 INSULIN PRECURSOR. [1] ; mRNA

sequence.

ACCESSION BQ272395

VERSION BQ272395.1 GI:20497464

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 174)

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,

Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,

Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,

Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,

Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R.,

Williams,T., Jackson,Y. and Bowers,Y.

Endocrine Pancreas Consortium

Unpublished (2000)

CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Other ESTs: iJ98c08.y1

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Alignment Scores:
Pred. No.: 5,18e-17 Length: 258
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-783-095-10 (1-26) x US-10-383-285-1 (1-258)

QY 1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGlySerLeuGlnLys 20
DB 133 GCGGTGGCCGCGCGCGGTCCCTCCAGCGCTCGAGGCTCCCTCCAGAAG 192
QY 21 ArgGlyIleValGluGln 26
DB 193 CGCGGTATCGTGGAGCAG 210

RESULT 2

US-09-807-742-16
; Sequence 16, Application US/09807742
; Publication No. US20030204864A1
; GENERAL INFORMATION:
; APPLICANT: DANIELL, HENRY
; TITLE OF INVENTION: PRODUCTION OF PHARMACEUTICAL PROTEINS IN TRANSGENIC
; FILE OF INVENTION: PLASTIDS
; FILE REFERENCE: 1465-PCT-US-00
; CURRENT APPLICATION NUMBER: US/09/807,742
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US01/06288
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 260
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-807-742-16

Alignment Scores:
Pred. No.: 5,21e-17 Length: 260
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-783-095-10 (1-26) x US-09-807-742-16 (1-260)

QY 1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGlySerLeuGlnLys 20
DB 133 GCGGGGGCCCTGGTGAGGAGCGCTCGAGCCCTTGGCCCTCGAGGGGTCCCTGCAGAAG 192
QY 21 ArgGlyIleValGluGln 26
DB 193 CGTGGCATTGTGAACAA 210

RESULT 3

US-09-807-742-17
; Sequence 17, Application US/09807742
; Publication No. US20030204864A1
; GENERAL INFORMATION:
; APPLICANT: DANIELL, HENRY
; TITLE OF INVENTION: PRODUCTION OF PHARMACEUTICAL PROTEINS IN TRANSGENIC
; FILE OF INVENTION: PLASTIDS
; FILE REFERENCE: 1465-PCT-US-00
; CURRENT APPLICATION NUMBER: US/09/807,742
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US01/06288
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 17
; LENGTH: 260
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chloroplast
; OTHER INFORMATION: modified proinsulin sequence
US-09-807-742-17

Alignment Scores:
Pred. No.: 5,21e-17 Length: 260
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-783-095-10 (1-26) x US-09-807-742-17 (1-260)

QY 1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGlySerLeuGlnLys 20
DB 133 GGTGGTGGTCTCTGGTGGTCTTTACACGCTTTAGCTTTAGAGGTTCTTTACAAAA 192
QY 21 ArgGlyIleValGluGln 26
DB 193 CGTGGTATGTAGAACAA 210

RESULT 4

US-10-746-149-23
; Sequence 23, Application US/10746149
; Publication No. US20040172667A1
; GENERAL INFORMATION:
; APPLICANT: Cooper, Richard K.
; APPLICANT: Fioretti, William C.
; APPLICANT: Cadd, Gary G.
; TITLE OF INVENTION: Administration of Transposon-Based Vectors to Reproductive Organs
; FILE REFERENCE: 51687-0280 (51687-294923)
; CURRENT APPLICATION NUMBER: US/10/746,149
; CURRENT FILING DATE: 2003-12-24
; PRIOR APPLICATION NUMBER: US 60/392,415
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/441,381
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,447
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,405
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,502
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,392
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,377
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 10/509,019
; PRIOR FILING DATE: 2003-08-26
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 260
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-746-149-23

Alignment Scores:
Pred. No.: 5,21e-17 Length: 260
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 22, 2004, 21:33:28 ; Search time 69 Seconds

(without alignments)
267.933 Million cell updates/sec

Title: US-10-783-095-10

Perfect score: 26

Sequence: 1 GGSPGAGSLPLALEGSLQKRGIVEQ 26

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1642189

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10783095/runat_22112004_143632_25659/app.query.fasta_1.199
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=P2Noligo.rni -MINMATCH=0.1
-LOOPEX=0 -LOOPEXT=0 -UNITS=bases -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPS=ZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10783095 @CGN 1 1 69 @runat_22112004_143632_25659
-NCPU=6 -ICPU=3 -NO.MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	258	2	US-08-600-783-17
2	26	100.0	269	2	Sequence 17, Appl
3	26	100.0	275	1	Sequence 10, Appl
4	26	100.0	275	1	Sequence 12, Appl
5	26	100.0	275	1	Sequence 14, Appl
6	26	100.0	275	1	Sequence 12, Appl
7	26	100.0	275	1	Sequence 14, Appl
8	26	100.0	276	1	Sequence 16, Appl
9	26	100.0	281	1	Sequence 15, Appl
10	26	100.0	281	1	Sequence 12, Appl
11	26	100.0	281	6	Sequence 13, Appl
12	26	100.0	290	1	Patent No. 5514646
					Sequence 41, Appl

c	13	26	100.0	290	1	US-08-030-731A-42	Sequence 42, Appl
	14	26	100.0	298	1	US-07-826-928A-28	Sequence 28, Appl
	15	26	100.0	304	1	US-07-698-551B-12	Sequence 12, Appl
	16	26	100.0	330	4	US-08-472-701-1	Sequence 1, Appl
	17	26	100.0	330	5	PCT-US95-08596-1	Sequence 1, Appl
	18	26	100.0	351	4	US-09-280-030-49	Sequence 49, Appl
	19	26	100.0	359	3	US-08-589-028-3	Sequence 3, Appl
	20	26	100.0	359	3	US-08-784-582-3	Sequence 3, Appl
	21	26	100.0	359	3	US-08-785-271-3	Sequence 3, Appl
	22	26	100.0	390	4	US-09-280-030-48	Sequence 48, Appl
	23	26	100.0	416	3	US-08-945-140-5	Sequence 5, Appl
	24	26	100.0	450	4	US-09-185-852-1	Sequence 1, Appl
	25	26	100.0	450	4	US-09-323-738-1	Sequence 1, Appl
	26	26	100.0	510	1	US-07-918-953-7	Sequence 7, Appl
	27	26	100.0	510	1	US-08-081-661-7	Sequence 7, Appl
	28	26	100.0	515	3	US-08-589-028-1	Sequence 1, Appl
	29	26	100.0	515	3	US-08-784-582-1	Sequence 1, Appl
	30	26	100.0	515	3	US-08-785-271-1	Sequence 1, Appl
	31	26	100.0	2499	4	US-09-775-508C-7	Sequence 7, Appl
	32	26	100.0	4992	4	US-09-015-399-10	Sequence 10, Appl
	33	21	80.8	87	4	US-09-676-787-1	Sequence 1, Appl
c	34	21	80.8	277	1	US-07-715-183C-3	Sequence 3, Appl
	35	21	80.8	278	1	US-07-715-183C-2	Sequence 2, Appl
	36	19	73.1	68	2	US-08-600-783-18	Sequence 18, Appl
	37	19	73.1	4646	4	US-09-485-286-14	Sequence 14, Appl
	38	17	65.4	99	1	US-07-918-953-3	Sequence 3, Appl
	39	17	65.4	99	1	US-08-081-661-3	Sequence 3, Appl
	40	17	65.4	598	3	US-09-537-696-11	Sequence 11, Appl
	41	17	65.4	633	3	US-09-537-696-12	Sequence 12, Appl
	42	17	65.4	666	3	US-09-537-696-13	Sequence 13, Appl
	43	16	61.5	51	6	5514646-42	Patent No. 5514646
c	44	15	57.7	74	4	US-09-485-286-9	Sequence 9, Appl
c	45	14	53.8	51	6	5514646-43	Patent No. 5514646

ALIGNMENTS

RESULT 1

US-08-600-783-17
; Sequence 17, Application US/08600783
; Patent No. 5962267
; GENERAL INFORMATION:
; APPLICANT: SHIN, Hang Cheol
; APPLICANT: CHANG, Seung Gu
; APPLICANT: KIM, Dae Young
; APPLICANT: KIM, Chong Suh
; TITLE OF INVENTION: Proinsulin Derivative and Process
; TITLE OF INVENTION: for Producing Human Insulin
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHIN, Hang Cheol
; STREET: Ssangma-Hanshin Apt. 102-1206,
; CITY: Kwangmyung-shi
; STATE: Kyungki-do
; COUNTRY: Republic of Korea
; ZIP: 423-030
; ADDRESSEE: CHANG, Seung Gu
; STREET: Hyundai Apt. 71-203, Apkujong-dong,
; CITY: Kangnam-ku
; STATE: Seoul
; COUNTRY: Republic of Korea
; ZIP: 135-110
; ADDRESSEE: KIM, Dae Young
; STREET: Sosa Jukong Apt. 108-202, Sosa Bon-dong,
; CITY: Sosa-ku
; STATE: Bucheon-shi
; COUNTRY: Republic of Korea
; ZIP: 422-230
; ADDRESSEE: KIM, Chong Suh
; STREET: Garden Heights Apt. 202-801, #100,

STREET: Hwangkeum-dong, Soosung-ku
CITY: Taegu
STATE: Taegu
COUNTRY: Republic of Korea
ZIP: 706-040
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk, 3.5 inch, 1.44MB storage
COMPUTER: IBM PC/AT
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,783
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: KR 95-2751
FILING DATE: 15-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Shahan Islam
REGISTRATION NUMBER: 32,507
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-1000
TELEFAX: (212) 953-7249
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 258 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-600-783-17

Alignment Scores:
Pred. No.: 1.9e-17 Length: 258
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-783-095-10 (1-26) x US-08-600-783-17 (1-258)

QY 1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGlySerLeuGlnLys 20
Db 133 GCGCGTGTCCGGGTGCGAGGCTCTGCGAGCGTGTGCGTGAAGGTTCCCTGCAGAA 192
QY 21 ArgGlyIleValGluGln 26
Db 193 CGTGGCATCGTTGACAA 210

RESULT 2
US-08-673-312-10
Sequence 10, Application US/08673312
Patent No. 5891699
GENERAL INFORMATION:
APPLICANT: BOULAIN, JEAN-CLAUDE
APPLICANT: CATTOLICO, LAURENCE
APPLICANT: DUCANCEL, FREDERIC
APPLICANT: MENEZ, ANDRE
TITLE OF INVENTION: MODIFIED BACTERIAL ALKALINE PHOSPHATASES
AND THEIR APPLICATION
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, WAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/673,312
FILING DATE: 28-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95-07833
FILING DATE: 29-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 846-380-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 269 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
US-08-673-312-10

Alignment Scores:
Pred. No.: 1.97e-17 Length: 269
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-783-095-10 (1-26) x US-08-673-312-10 (1-269)

QY 1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGlySerLeuGlnLys 20
Db 138 GCGCGGGGCCCGCGCGCGAGCCTGCACCGCTGGCGCTGGAGGGCAGCCTCCAGAA 197
QY 21 ArgGlyIleValGluGln 26
Db 198 CGTGGCATTTGGACAG 215

RESULT 3
US-07-918-953-12
Sequence 12, Application US/07918953
Patent No. 5268453
GENERAL INFORMATION:
APPLICANT: ANDY, ROBIN J.
APPLICANT: LARSON, ERIC R.
TITLE OF INVENTION: TISSUE-SELECTIVE INSULIN ANALOGS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/918,953
FILING DATE: 19920730
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BOZICEVIC, KARL
REGISTRATION NUMBER: 28,807
REFERENCE/DOCKET NUMBER: 21900-2027420

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 22, 2004, 20:48:38 ; Search time 327 Seconds

(without alignments)
417.385 Million cell updates/sec

Title: US-10-783-095-10

Perfect score: 26

Sequence: 1 GGCGAGSLQPLAESLQKRGIVEQ 26

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8264066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_17/USPRO_spool/US10783095/runat_22112004_143630_25620/app_query.fasta_1.199
-DB=N Geneseq 23Sep04 -QMT=fastap -SUFFIX=P2Noligo.rng -MINWATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCA=IGN=200 -THR SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10783095 @CGN 1 1 470 @runat_22112004_143630_25620
-NCPU=6 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N Geneseq 23Sep04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	258	2 AAT37134	Aat37134 Proinsuli
2	26	100.0	258	10 ADCS1568	Adcs1568 Human pro
3	26	100.0	258	10 ADf16306	Adf16306 Human alb
4	26	100.0	258	10 ADH21708	Adh21708 Human lon
5	26	100.0	260	12 ADI04119	Adi04119 Proinsuli
6	26	100.0	261	2 AAT17831	Aat17831 Optimised

7	26	100.0	261	2 AAT17830	Aat17830 Human ins
8	26	100.0	268	2 AAQ43935	Aac43935 MetTyr hu
9	26	100.0	269	2 AAT80401	Aat80401 Proinsuli
10	26	100.0	272	1 AAN20041	Aan20041 Human pro
11	26	100.0	272	1 AAN60872	Aan60872 Sequence
12	26	100.0	275	2 AAQ37003	Aaq37003 Proinsuli
13	26	100.0	279	2 AAQ83195	Aaq83195 Proinsuli
14	26	100.0	281	2 AAQ05663	Aaq05663 Human pro
15	26	100.0	281	2 AAQ38310	Aaq38310 hpi gene
16	26	100.0	283	2 AAQ20543	Aaq20543 Encodes p
17	26	100.0	289	12 ADL24436	Adl24436 Modified
18	26	100.0	290	12 ADL24441	Adl24441 Modified
19	26	100.0	298	2 AAQ27123	Aaq27123 DNA encod
20	26	100.0	298	2 AAQ99025	Aaq99025 Chelating
21	26	100.0	304	2 AAQ15230	Aaq15230 Ncol-Sali
22	26	100.0	327	6 AEN85338	Abn85338 Oligonuc
23	26	100.0	330	9 ADA09217	Ada09217 Human mRN
24	26	100.0	330	12 ADJ25929	Adj25929 Human pre
25	26	100.0	333	3 AAZ51476	Aaz51476 Human ins
26	26	100.0	342	1 AAN40179	Aan40179 Sequence
27	26	100.0	342	1 AAN50152	Aan50152 Sequence
28	26	100.0	342	1 AAN50082	Aan50082 Proinsuli
29	26	100.0	351	3 AAZ59208	Aaz59208 WWPsp-WMP
30	26	100.0	359	2 AAT75193	Aat75193 Human ins
31	26	100.0	359	2 AAT75652	Aat75652 Human ins
32	26	100.0	359	3 AAC55717	Aac55717 Mutant hu
33	26	100.0	359	4 AAF58802	Aaf58802 Human ins
34	26	100.0	390	3 AAZ59207	Aaz59207 WWPsp-WMP
35	26	100.0	416	1 AAN10052	Aan10052 mRNA sequ
36	26	100.0	416	1 AAN40256	Aan40256 Sequence
37	26	100.0	450	3 AAZ29531	Aaz29531 Human pro
38	26	100.0	450	4 AAD17486	Aad17486 Human pre
39	26	100.0	450	6 ABR81193	Abk81193 cDNA enco
40	26	100.0	450	10 ACC78888	Acc78888 Humanised
41	26	100.0	450	12 ADL24432	Adl24432 Human ins
42	26	100.0	450	12 ADN49715	Adn49715 Human ins
43	26	100.0	462	3 AA440597	Aa440597 Human pre
44	26	100.0	496	3 AAA95514	Aa95514 Insert ir
45	26	100.0	515	3 AAC55716	Aac55716 Human ins

ALIGNMENTS

RESULT 1

AAT37134
ID AAT37134 standard; DNA; 258 BP.
XX
XX
AC AAT37134;
XX
XX 15-APR-1997 (first entry)
XX
XX Proinsulin coding sequence.
XX
XX Proinsulin; human; beta-turn peptide; insulin; A chain; protein folding;
KW enzymatic cleavage; B chain; miniproinsulin; enzymatic hydrolysis; ss.
XX
XX Synthetic.
XX
XX GB2298206-A.
XX
XX 28-AUG-1996.
XX
XX 14-FEB-1996; 96GB-00002998.
XX
XX 15-FEB-1995; 95KR-00002751.
XX
XX (HANI-) HANIL SYNTHETIC FIBER CO LTD.
XX
XX Shin H, Chang S, Kim D, Kim C;
XX
XX WPI; 1996-373460/38.
XX
XX New human pro-insulin derivs. - comprising A and B chains linked by small

PT

PT beta-turn peptide.
 PS Example 4; Page 18; 52pp; English.
 XX
 CC This sequence represents a synthetic human proinsulin coding sequence.
 CC The codon preference of this sequence has been altered to favour the
 CC preferred codons in *E. coli*. This sequence can be used in the human
 CC proinsulin derivatives of the invention. In the derivatives of the
 CC invention, a beta-turn peptide (see AAW03935-W03938) is used to join the
 CC human insulin A and B chains. Beta-turn peptides are thought to be sites
 CC for initiation of protein folding. This is thought to be due to the fact
 CC that they are determined by short-range interactions, and therefore limit
 CC the conformational space available to the polypeptide chain. Beta-turns
 CC also play a valuable role in relation to enzymatic cleavage. By using the
 CC insulin derivatives (also referred to as miniproinsulin), refolding and
 CC hydrolysing processes can be carried out more efficiently than with
 CC proinsulin (which contains a bulky C-peptide). The derivatives can be
 CC recovered from transformed host cells in higher refolding yields than
 CC human proinsulin. The derivatives are also readily converted to insulin
 CC by enzymatic hydrolysis, such as with trypsin and carboxypeptidase B. The
 CC insulin derivatives are easier to produce than current derivatives. By
 CC using the derivatives of the invention, insulin can be produced at higher
 CC yields than with current techniques
 XX
 SQ Sequence 258 BP; 50 A; 67 C; 75 G; 66 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.52e-15 Length: 258
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-10-783-095-10 (1-26) x AAT37134 (1-258)
 QY 1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGlySerLeuGlnLys 20
 DB 133 GCGGGTGGTCGGGTGCGAGGCTCTCTGCAGCGCTTGGCGTGGAGAGTTCCCTGCAGAAA 192
 QY 21 ArgGlyIleValGluGln 26
 DB 193 CGTGGCATCGTTGAACAA 210
 RESULT 2
 ADCS1568
 ID ADCS1568 standard; DNA; 258 BP.
 XX
 AC ADCS1568;
 XX
 DT 18-DEC-2003 (first entry)
 DE Human proinsulin-encoding nucleotide sequence #SEQ ID 1.
 XX
 KW Antidiabetic; insulin; diabetes; genetic engineering; plant tissue; ds;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003076595-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 07-MAR-2003; 2003WO-US007316.
 XX
 PR 08-MAR-2002; 2002US-0362874P.
 XX
 PA (PROD-) PRODIGENE INC.
 XX
 PI Streatfield S, Howard JA;
 XX
 DR WPI; 2003-748386/70.
 XX

PT New plant or plant cells expressing an insulin or an insulin-like
 PT protein, useful as sources of (pro)insulin, particularly for producing
 PT (pro)insulin proteins in commercial quantities, or as oral replacement
 PT for insulin injections.
 XX
 XX Example 1; SEQ ID NO 1; 26pp; English.
 XX
 CC The invention relates to a plant, which expresses an insulin or an
 CC insulin-like protein. The plant or plant cells are useful as sources of
 CC (pro)insulin for a variety of purposes, particularly for producing
 CC (pro)insulin proteins in commercial quantities. Plant tissue can be
 CC orally administered to diabetic animals as an alternative to insulin
 CC injections. The plant-produced proteins are also useful for providing
 CC less expensive and more readily available source of the protein as
 CC reagent, or in other experimentation involving insulin proteins. The
 CC current sequence represents the human proinsulin-encoding nucleotide
 CC sequence used in an example from the invention for introduction into
 CC maize seed.
 XX
 SQ Sequence 258 BP; 38 A; 98 C; 78 G; 44 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.52e-15 Length: 258
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-10-783-095-10 (1-26) x ADCS1568 (1-258)
 QY 1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGlySerLeuGlnLys 20
 DB 133 GCGGGTGGCGCGCGCGCGCTCTCTGCAGCGCTTCCCGCTCCCTCCAGAG 192
 QY 21 ArgGlyIleValGluGln 26
 DB 193 CGCGGTATCGTGGAGCAG 210
 RESULT 3
 ADF16306
 ID ADF16306 standard; DNA; 258 BP.
 XX
 AC ADF16306;
 XX
 DT 12-FEB-2004 (first entry)
 DE Human albumin fusion protein-related DNA sequence SeqID1398.
 XX
 KW albumin fusion protein; albumin activity; human serum albumin;
 KW serum osmotic pressure; shelf-life; stability; antidiabetic;
 KW gene therapy; diabetes mellitus; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003060071-A2.
 XX
 PD 24-JUL-2003.
 XX
 PF 23-DEC-2002; 2002WO-US040891.
 XX
 PR 21-DEC-2001; 2001US-0341811P.
 PR 24-JAN-2002; 2002US-0350358P.
 PR 28-JAN-2002; 2002US-0351360P.
 PR 26-FEB-2002; 2002US-0359370P.
 PR 28-FEB-2002; 2002US-0360000P.
 PR 27-MAR-2002; 2002US-0367500P.
 PR 08-APR-2002; 2002US-0370227P.
 PR 10-MAY-2002; 2002US-0378950P.
 PR 24-MAY-2002; 2002US-0382617P.
 PR 28-MAY-2002; 2002US-0383123P.
 PR 05-JUN-2002; 2002US-0385708P.
 PR 10-JUL-2002; 2002US-0394625P.
 PR

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 22, 2004, 20:49:53 ; Search time 2542 Seconds
(without alignments)
483.687 Million cell updates/sec

Title: US-10-783-095-10
Perfect score: 26
Sequence: 1 GGGPGAGSLQPLAEGSLQRRGIVBQ 26

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Word size: 1

Total number of hits satisfying chosen parameters: 9044172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10783095/runat_22112004_143630_25630/app_query.fasta_1.199
-DB=GenEmbl -QMT=fastap -SUFFIX=P2Noligo.rge -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-USER=US10783095 -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-OUTER=US10783095 -CGEN 1 1 3731 @runat_22112004_143630_25630 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=130 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb.ba.*
2: gb.hg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	258	AR077833	AR077833 Sequence
2	26	100.0	261	A48809	A48809 Sequence 1
3	26	100.0	261	A48810	A48810 Sequence 2
4	26	100.0	269	AR069914	AR069914 Sequence

c	5	26	100.0	271	6	E43909	E43909 Novel vecto
6	7	26	100.0	272	6	I02917	I02917 Sequence 6
7	8	26	100.0	273	6	E00074	E00074 DNA sequence
8	9	26	100.0	273	6	I00910	I00910 Sequence 2
9	10	26	100.0	275	6	I14159	I14159 Sequence 12
10	11	26	100.0	275	6	I14160	I14160 Sequence 14
11	12	26	100.0	276	6	I14161	I14161 Sequence 16
12	13	26	100.0	277	12	ARHSINS	V00082 Artificial
13	14	26	100.0	279	6	E43908	E43908 Novel vecto
14	15	26	100.0	286	12	SYNINSGSA	M12913 Synthetic h
15	16	26	100.0	289	6	I03833	I03833 Sequence 35
16	17	26	100.0	290	6	A07328	A07328 Synthetic D
17	18	26	100.0	290	6	A07329	A07329 Synthetic D
18	19	26	100.0	290	6	A11939	A11939 Nucleotide
19	20	26	100.0	290	6	A11940	A11940 Nucleotide
20	21	26	100.0	290	6	I12514	I12514 Sequence 41
21	22	26	100.0	298	6	I13722	I13722 Sequence 28
22	23	26	100.0	330	6	AR275652	AR275652 Sequence
23	24	26	100.0	333	6	BD224077	BD224077 Method fo
24	25	26	100.0	333	9	BT006808	BT006808 Homo sapi
25	26	26	100.0	333	12	BT007778	BT007778 Synthetic
26	27	26	100.0	342	6	I00596	I00596 Sequence 4
27	28	26	100.0	342	6	I02324	I02324 Sequence 4
28	29	26	100.0	342	6	I03602	I03602 Sequence 4
29	30	26	100.0	342	6	I07905	I07905 Sequence 15
30	31	26	100.0	346	6	E00054	E00054 DNA encodin
31	32	26	100.0	346	6	E00191	E00191 DNA coding
32	33	26	100.0	346	6	E00575	E00575 Synthetic D
33	34	26	100.0	351	6	E39198	E39198 DNA encodin
34	35	26	100.0	351	6	AR274665	AR274665 Sequence
35	36	26	100.0	351	12	SYNINSGS	J02547 Human (synt
36	37	26	100.0	359	6	AR102882	AR102882 Sequence
37	38	26	100.0	359	6	AR108061	AR108061 Sequence
38	39	26	100.0	359	6	AR134709	AR134709 Sequence
39	40	26	100.0	390	6	E39197	E39197 DNA encodin
40	41	26	100.0	390	6	AR274664	AR274664 Sequence
41	42	26	100.0	392	9	MAGINS	J00336 Monkey (M.
42	43	26	100.0	416	6	A57719	A57719 Sequence 5
43	44	26	100.0	416	6	AR175913	AR175913 Sequence
44	45	26	100.0	416	6	E00011	E00011 RNA coding

ALIGNMENTS

RESULT 1	AR077833	Sequence 17 from patent US 5962267.	258 bp	DNA	linear	PAT 31-AUG-2000
LOCUS	AR077833	Sequence 17 from patent US 5962267.	258 bp	DNA	linear	PAT 31-AUG-2000
DEFINITION	AR077833	Sequence 17 from patent US 5962267.	258 bp	DNA	linear	PAT 31-AUG-2000
ACCESSION	AR077833	Sequence 17 from patent US 5962267.	258 bp	DNA	linear	PAT 31-AUG-2000
VERSION	AR077833.1	GI:10004579	258 bp	DNA	linear	PAT 31-AUG-2000
KEYWORDS	Unknown.	Unknown.	258 bp	DNA	linear	PAT 31-AUG-2000
SOURCE	Unknown.	Unknown.	258 bp	DNA	linear	PAT 31-AUG-2000
ORGANISM	Unknown.	Unknown.	258 bp	DNA	linear	PAT 31-AUG-2000
REFERENCE	1 (bases 1 to 258)	Shin,H.-C., Chang,S.-G., Kim,D.-Y. and Kim,C.-S.	258 bp	DNA	linear	PAT 31-AUG-2000
AUTHORS	Shin,H.-C., Chang,S.-G., Kim,D.-Y. and Kim,C.-S.	Proinsulin derivative and process for producing human insulin	258 bp	DNA	linear	PAT 31-AUG-2000
TITLE	Proinsulin derivative and process for producing human insulin	Proinsulin derivative and process for producing human insulin	258 bp	DNA	linear	PAT 31-AUG-2000
JOURNAL	Patent: US 5962267-A 17 05-OCT-1999;	Patent: US 5962267-A 17 05-OCT-1999;	258 bp	DNA	linear	PAT 31-AUG-2000
FEATURES	Location/Qualifiers	Location/Qualifiers	258 bp	DNA	linear	PAT 31-AUG-2000
source	1..258	/organism="unknown"	258 bp	DNA	linear	PAT 31-AUG-2000
source	/mol_type="unassigned DNA"	/mol_type="unassigned DNA"	258 bp	DNA	linear	PAT 31-AUG-2000

Alignment Scores:	1.13e-17	Length:	258
Pred. No.:	26.00	Matches:	26
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	6		

US-10-783-095-10 (1-26) x AR077833 (1-258)

1	GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGluGlySerLeuGlnI	20
Qy		
133	GGCGGTGTCGGGGTGGAGGCTCTCTGAGCGGTGGCGTGGAGGTTCCTTCGAGAA	192
DB		
21	ArgGlyIleValGluGln	26
Qy		
193	CGTGGCATCGTTGAACAA	210
DB		

RESULT 2	LOCUS	DEFINITION	SEQUENCE	261 bp	DNA	linear	PAT 07-MAR-1997
A48809	LOCUS	Sequence 1 from Patent EP0704527.					
A48809	DEFINITION	Sequence 1 from Patent EP0704527.					
A48809	ACCESSION	A48809					
A48809	VERSION	A48809.1	GI:2302472				
A48809	KEYWORDS	.					
A48809	SOURCE	unidentified					
A48809	ORGANISM	unidentified					
A48809	ORGANISM	unclassified					

REFERENCE AUTHORS TITLE JOURNAL COMMENT	1
Meatrick, S., Punt, P.J., Valinger R., Van and Den.H.C. DNA sequences encoding biosynthetic insulin precursors and process for preparation of insulin	
Patent: EP 0704527-A 1 03-APR-1996; PLIVA PHARM & CHEM WORKS (YU)	
Other publication CN 1126761 960717	
Other publication CA 2155451 960206	
Other publication SK 97195 960207	
Other publication SI 9500250 960229	
Other publication BG 99844 960229	
Other publication CZ 9501999 960214	
Other publication PL 309882 960219.	

```

ORIGIN
Alignment Scores:
Pred. No.:      1.14e-17
Score:          26.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:    100.00%
DB:             6
Length:         261
Matches:        26
Conservative:   0
Mismatch:       0
Indels:         0
Gaps:           0
US-10-783-095-10 (1-26) x A48809 (1-261)

```

1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGluGlySerLeuGlnLys 20
133 GCGGGGGCCCTGGTGCAGGGCAGCCTGAGCCCTGGCCCTGAGGGGTCCCTGCAGAAG 192
21 ArgGlyIleValGluGln 26
193 CGTGGCATTTGTGGACAA 210

RESULT 3	LOCUS	DEFINITION	SEQUENCE 2 FROM PATENT	261 bp	DNA	linear	PAT 07-MAR-1997
A48810	A48810	SEQUENCE 2 FROM PATENT	EP0704527.				
ACCESSION	A48810						
VERSION	A48810.1	GI:2302473					
KEYWORDS	.						
SOURCE	unidentified						
ORGANISM	unidentified						
	unclassified						

REFERENCE	AUTHORS	TITLE	JOURNAL
1	Meetric, S., Punt, P.J., Valinger, R., Van and Den, H.C.	DNA sequences encoding biosynthetic insulin precursors and process for preparation of insulin	Patent: EP 0704527-A 2 03-APR-1996;
			unclassified.

```

PLIVA PHARM & CHEM WORKS (YU)
Other publication CN 1126761 960717
Other publication CA 2155431 960206
Other publication SK 97195 960207
Other publication SI 9500250 960229
Other publication BG 95844 960223
Other publication CZ 9501999 960214
Other publication PL 309882 960219.
Location/Qualifiers
1. .361
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1. .358
/ncbi="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"
/ccdon_start=1
/protein_id="CAA03148.1"
/db_xref="GI:2302474"
/translation="EVNGLCGSHLVEALVLVGCGFFVTPTKTRAEADLQVGQVEL
CCGACRACRCHPAPGCSQKGGKGGCCGCTGCSLYOLENVCN."

```

[illegible]

1 GlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGluGlySerLeuGlnLys 20
QY